

University of Groningen

Quantitative approaches for profit maximization in direct marketing

van der Scheer, H.R.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

1998

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

van der Scheer, H. R. (1998). *Quantitative approaches for profit maximization in direct marketing*. [Thesis fully internal (DIV), University of Groningen]. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Samenvatting

Het succes van een *direct-mail*campagne wordt bepaald door (1) de selectie van meest kansrijke adressen, de zogenaamde *targets*, (2) het aanbod, (3) het design van de mailing en (4) het tijdstip waarop en de frequentie waarmee de mailing wordt verstuurd. De selectie van de targets wordt zowel in theorie als in de praktijk als het belangrijkste element beschouwd. Een aanzienlijke gedeelte van direct-marketingonderzoek richt zich dan ook op target selectietechnieken. Recentelijk is in de literatuur ook aandacht besteed aan het manipuleren van het design van de mailing. Het doel van dit proefschrift is om methoden te introduceren waarmee de effectiviteit, en daarmee de winst, van een direct-mailcampagne kan worden verhoogd.

In hoofdstuk 2 wordt een overzicht gegeven van de technieken die worden gebruikt om targets te selecteren. Helaas is het niet mogelijk om een eenduidige uitspraak te doen over wat de beste techniek is. De reden hiervoor is dat verschillende, soms conflicterende, aspecten een rol spelen bij de beoordeling van een model. Daarnaast is er nauwelijks literatuur die de technieken vergelijkt op basis van de verwachte winst. Drie algemene conclusies kunnen worden getrokken ten aanzien van deze technieken. Ten eerste, exploratieve methoden zoals CHAID zijn minder geschikt om targets te selecteren dan regressiemethoden (probit, discriminant analyse). Ten tweede, regressiemethoden werken goed en hebben als voordeel dat ze eenvoudig te interpreteren zijn. Ten derde, de resultaten van neurale netwerken zijn niet beter dan die van de conventionele statistische technieken.

De meeste selectietechnieken richten zich op mailings waarbij respons een vaste opbrengst oplevert. De technieken beschouwen dus alleen de binaire keuze van het al dan niet responderen. Echter, veel direct-mailcampagnes generen niet een simpele binaire respons maar een respons waarbij de opbrengst, oftewel het bestede bedrag, varieert over de individuen. Het bestede bedrag is bijvoorbeeld de totale besteding bij een postorderbedrijf of de donatie aan een charitatieve instelling. In hoofdstuk 3 specificeren we een model dat zowel de

binaire respons (p) als het bestede bedrag (a) modelleert. We leiden een selectieregel af die gebaseerd is op p en a ; deze regel maximaliseert de verwachte winst. Om deze selectieregel te kunnen toepassen moeten de dichtheden van het geschatte bestede bedrag van de respondenten en de non-respondenten in beschouwing worden genomen. We geven drie benaderingen om deze dichtheden te bepalen. De eerste benadering verwaarloost simpelweg het verschil in de twee dichtheden. De selectieregel is dan een intuïtief aantrekkelijke curve in de (p, a) -ruimte. De tweede benadering gaat er vanuit dat beide dichtheden normaal zijn met een verschillend gemiddelde en dezelfde variantie. Ook deze benadering resulteert in een selectieregel die weergegeven kan worden in een eenvoudige curve in de (p, a) -ruimte. De derde benadering schat de dichtheden met een non-parametrische schattingsmethode. Dit geeft een wat complexere curve in de (p, a) -ruimte. De resultaten van de toepassing laten zien dat het belangrijk is om het bestede bedrag te modelleren. Zelfs de meest eenvoudige benadering genereert aanzienlijk meer winst dan wanneer alleen binaire respons wordt gemodelleerd.

De parameters van een target selectiemodel zijn meestal onbekend en moeten worden geschat. Op basis van deze schattingen worden individuen geselecteerd. Op deze manier wordt het schatten van de parameters en het selecteren van de individuen in twee aparte stappen uitgevoerd. Aangezien door het apart beschouwen van deze twee stappen de schattingsonzekerheid wordt genegeerd, leidt dit tot een sub-optimale selectieregel en daardoor niet tot de optimale verwachte winst. In hoofdstuk 4 integreren we deze twee stappen. Dit resulteert in een Bayesiaanse selectieregel die volgt uit de winstfunctie van de organisatie. Een van de moeilijkheden van deze aanpak is het evalueren van de meerdimensionale integraal die deze Bayesiaanse selectieregel definieert. We bespreken drie methoden om deze integraal te evalueren. Bij de eerste methode nemen we aan dat de *posterior* verdeling normaal verdeeld is. In dat geval kan de integraal worden geschreven als een gesloten uitdrukking en kan dan eenvoudig worden berekend. Een nadeel van deze methode is dat eventuele priorinformatie wordt verwaarloosd. Een verfijnde methode om de integraal te evalueren is de *Laplace* benadering. De berekeningen bij deze benadering bestaan uit differentiëren en maximaliseren (één Newton-Raphson stap). De integraal hoeft dan dus niet numeriek te worden bepaald. De derde benadering is met *Markov chain Monte Carlo* integratie. Deze methode is, met name door de opkomst van krachtige personal computers, de laatste jaren populair geworden. Ook deze methode is relatief eenvoudig te implementeren. We hebben de tweede en derde methode toegepast met zowel een informatieve

prior als een niet-informatieve prior. De resultaten laten zien dat de Bayesiaanse selectieregel inderdaad meer winst genereert; de verschillen tussen de drie benaderingen zijn relatief klein.

De kwaliteit van de *mailinglist*, de database met informatie over alle potentiële targets, is een belangrijk aspect voor het succes van een campagne. De waarde van de informatie op de mailinglist is de mate waarin de targets kunnen worden geïdentificeerd. In hoofdstuk 5 bekijken we de relatie tussen de kwaliteit van informatie en winst. Eerst analyseren we de relatie tussen de R^2 , een indicator voor de kwaliteit, en de winst. Zoals verwacht neemt de winst toe met R^2 . Daarna beschouwen we de waarde van postcode-informatie. Deze informatie comprimeert, in Nederland, de informatie van gemiddeld 16 huishoudens. Een organisatie kan in plaats van informatie op individueel niveau postcode-informatie gebruiken om targets te selecteren. Hierdoor zal de maximale winst afnemen. De meerwaarde van individuele informatie ten opzichte van postcode-informatie hangt af van de mate van homogeniteit binnen de postcodegebieden. Indien de mate van homogeniteit groot is zal individuele informatie de selectieresultaten nauwelijks verbeteren. Als er nauwelijks sprake is van homogeniteit binnen een postcodegebied zal de selectie op basis van individuele informatie enorm verbeteren. We hebben de relatie afgeleid tussen de kwaliteit van informatie, uitgedrukt in de *intra-class* correlatie (een indicator voor homogeniteit), en de verwachte winst. Deze relatie stelt een organisatie in staat om de waarde van informatie vast te stellen. Dit is een belangrijke aspect voor een organisatie die keuze heeft tussen het aanschaffen van informatie op individueel niveau, postcode niveau of helemaal geen informatie. Alhoewel onze aanpak gestileerd en hoofdzakelijk theoretisch van aard is, laat het zien welke aspecten een rol spelen bij het waarderen van informatie.

Hoofdstuk 6 gaat over het manipuleren van de karakteristieken van de mailing. De traditionele aanpak om het effect van meerdere karakteristieken te analyseren is door elk karakteristiek apart te beschouwen. Meestal is zo een methode inefficiënt. We bespreken twee methoden, beide gebaseerd op conjunct analyse, om een optimaal design op een efficiëntere manier te bepalen. We illustreren dit met drie toepassingen. Beide methoden kunnen worden uitgebreid door de interactie tussen de karakteristieken van de targets en de karakteristieken van de mailing mee te nemen. Het is namelijk zeer aannemelijk dat deze interactie bestaat. Bijvoorbeeld, een direct marketing organisatie zal andere communicatie-elementen willen gebruiken voor targets die al lange tijd klant zijn dan voor targets die nog nooit iets bij de organisatie hebben gekocht. We geven een methode om simultaan het maildesign te optimaliseren

en de targets te selecteren. De toepassing laat zien dat er inderdaad sprake is van interactie; de optimale strategie resulteert in dertien verschillende soorten mailings. Elke target ontvangt een van deze mailings. Door deze strategie neemt de verwachte winst toe met 25% ten opzichte van de traditionele mailing en met 9% als één optimale mailing voor alle targets wordt gebruikt.

Marketing is in principe een continu proces. Een aankoop is in het algemeen niet het einde van een relatie tussen een individu en de organisatie maar het begin of een continuering. Het succes van een mailcampagne zou gedefinieerd moeten worden op basis van een langetermijncriterium in plaats van op een kortetermijncriterium. Een van de grote uitdagingen in direct marketing is om de bestaande methoden uit te breiden naar een langetermijnstrategie. Deze strategie zou zich moeten richten op het maximaliseren van de *lifetime value* van een individu. In het ideale geval wordt deze berekend op basis van individuele gegevens en wordt vervolgens gebruikt voor een normatieve strategie ten aanzien van het creëren, ontwikkelen en in stand houden van relaties. Het is duidelijk dat dit moeilijk te realiseren is. Echter, er zijn verscheidene mogelijkheden om bepaalde aspecten van zo'n strategie te verbeteren. In hoofdstuk 7 laten we dit zien door te bepalen met welke frequentie een organisatie mailings moet versturen. Het onderliggende idee van onze methode is dat de mailingfrequentie gebaseerd moet zijn op het aankoopgedrag van de consument. We specificeren een model dat het aankoopgedrag van de consument beschrijft. Dit model ontleedt het gedrag in de timing van de aankoop, die bestaat uit tussen aankooptijd en aankoopversnelling, en de keuze om het product via direct marketing te kopen of in een winkel. Dit model wordt gebruikt om de beslissingen van de consumenten te simuleren voor een range van mailingfrequenties. De uitkomsten van deze simulaties geven aan wat de maximale frequentie is. Een illustratief voorbeeld laat zien dat dit een relatief eenvoudige manier is om de optimale frequentie te bepalen.